ISSN: 2508-7894
KJAI website: http://accesson.kr/kjai
doi: http://doi.org/10.24225/kjai.2024.12.3.1

A Study on Improving License Plate Recognition Performance Using Super-Resolution Techniques

Kyeongseok JANG[1], Kwangchul SON[2]

Received: July 29, 2024. Revised: August 22, 2024. Accepted: September 05, 2024.

Abstract

In this paper, we propose an innovative super-resolution technique to address the issue of reduced accuracy in license plate recognition caused by low-resolution images. Conventional vehicle license plate recognition systems have relied on images obtained from fixed surveillance cameras for traffic detection to perform vehicle detection, tracking, and license plate recognition. However, during this process, image quality degradation occurred due to the physical distance between the camera and the vehicle, vehicle movement, and external environmental factors such as weather and lighting conditions. In particular, the acquisition of low-resolution images due to camera performance limitations has been a major cause of significantly reduced accuracy in license plate recognition. To solve this problem, we propose a Single Image Super-Resolution (SISR) model with a parallel structure that combines Multi-Scale and Attention Mechanism. This model is capable of effectively extracting features at various scales and focusing on important areas. Specifically, it generates feature maps of various sizes through a multi-branch structure and emphasizes the key features of license plates using an Attention Mechanism. Experimental results show that the proposed model demonstrates significantly improved recognition accuracy compared to existing vehicle license plate super-resolution methods using Bicubic Interpolation.

Keywords : Deep Learning, Single Image Super-Resolution, Multi-Scale, Attention Mechanism, License Plate Eecognition

Major Classification Code : Artificial Intelligence, Deep Learning, Super-Resolution, License Plate Recognition, etc

1. Introduction

In recent years, there has been a surge in research utilizing existing surveillance system imagery to enhance Intelligent Transportation Systems (ITS) (Bugeja et al., 2020; Qiu et al., 2024). This research spans various applications, including illegal parking enforcement, traffic monitoring, vehicle tracking, and traffic flow analysis using surveillance cameras (Dui et al., 2024; Ngeni et al., 2022; Hassan et al., 2021). These traffic detection systems typically involve vehicle detection, tracking, and license plate recognition from fixed camera footage (Qiu et al., 2021; Guerrero-Ibañez et al., 2021). However, real-world environments present numerous challenges that degrade image quality,

* This work was conducted during the sabbatical leave with support from Kwangwoon University in 2023.

1 First Author. Student, Department of Plasma Bio Display, Kwangwoon University, Republic of Korea. Email: ksjang1234@kw.ac.kr

2 Second Author. Professor, Department of Smart Electrical and Electronic Engineering, Kwangwoon University, Republic of Korea, Email: kcson@kw.ac.kr

such as the distance between cameras and vehicles, vehicle movement, weather variations, and lighting conditions. These factors directly impact the accuracy of license plate recognition. When performing recognition of characters and numbers within license plates, the occurrence of noise and loss of detailed information makes it difficult to distinguish between characters and numbers with similar shapes (for example, '8' and '0', '아' and '하' in Korean). Moreover, accurate feature extraction becomes challenging when using Convolution operations, leading to decreased recognition accuracy. In particular, Korean vehicle license plates pose additional difficulties in recognition due to their unique design and the use of Hangul characters. These issues can potentially lead to overall performance degradation in traffic management systems. To address these issues, researchers are increasingly turning to Deep Learning techniques, focusing on Super-Resolution (SR) technology (Yang et al., 2010; Dong et al., 2014; Saharia et al., 2022). SR aims to reconstruct high-resolution images from low-resolution inputs and is categorized into Multi Image Super-Resolution (MISR) and Single Image Super-Resolution (SISR) (Ledig et al., 2017; Salvetti et al., 2020). SISR is more suitable for real-time traffic monitoring systems, with Convolutional Neural Network (CNN)-based approaches demonstrating superior performance compared to traditional methods.

In this paper, we propose a SISR method incorporating Multi-Scale and Attention Mechanisms to improve license plate recognition accuracy. The Multi-Scale approach effectively extracts features of various sizes, while the Attention Mechanism emphasizes crucial features for license plate recognition. This method not only enhances resolution but also effectively addresses issues such as blurring, loss, and distortion. By emphasizing key features necessary for license plate recognition, this approach enables more effective resolution enhancement. Our proposed method aims to significantly improve the quality of traffic surveillance imagery, thereby enhancing the accuracy and reliability of license plate recognition systems in challenging real-world conditions.

2. Related Work

2.1. Character and Digit Recognition

To recognize characters and numbers in images and videos, computers require a conversion process to transform visual data into a comprehensible format. Traditionally, Optical Character Recognition (OCR) has been employed for this purpose (Arica & Yarman-Vural, 2002). OCR is a technology that extracts and recognizes characters and numbers from images or videos, converting them into a text format that computers can process. The OCR process begins with preprocessing, which involves adjusting the skew, resizing, and removing noise from text images or Portable Document Format (PDF) files. Following this, feature extraction takes place, where character or number contours are extracted and patterns are identified. The final stage involves character and number identification based on the extracted features, matching them to the most similar known characters (Memon et al., 2020). As artificial intelligence has progressed, various machine learning algorithms have been applied to OCR, particularly Support Vector Machines (SVM) and k-Nearest Neighbors (k-NN). Recent advancements in Deep Learning have led to the development of character and number recognition models based on Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) (Drobac & Lindén 2020; El Bahi & Zatni 2019). CNN-based models automatically learn and extract features from images, overcoming the limitations of traditional methods that require manual feature design. RNN-based models, especially those utilizing Long Short-Term Memory (LSTM), excel at sequential text recognition, proving particularly effective for handwritten or continuous text.

In this paper, we focus on Korean license plates, which have distinct characteristics. They feature white backgrounds with black characters and numbers. New plate formats have expanded to include three-digit numbers at the front and incorporate anti-forgery holograms. There are notable differences in font and size between old and new plates. The unique structure and combination method of Korean characters necessitate different recognition approaches compared to English-language plates. For instance, distinguishing between similar characters like '아' and '오' or '호' and '무' requires special attention in Korean license plate recognition. Given these unique features, CNN-based OCR methods demonstrate superior performance compared to traditional OCR approaches for Korean license plate recognition. CNNs can learn various plate features, enabling more robust recognition despite variations in font or size. Furthermore, data augmentation techniques can be employed to enhance recognition performance under diverse lighting conditions and angles.

2.2. Attention Mechanism

The Attention Mechanism was initially proposed to emphasize local contextual information in Recurrent Neural Network (RNN) based models. However, it is now widely used in Convolutional Neural Network (CNN) based models to highlight key features in images (Lieskovská et al., 2021; Chen et al., 2020; Cui et al., 2021; Cai & Wei, 2020). This mechanism, inspired by human visual attention capabilities, selectively focuses on specific parts of input data to improve model performance. Features in images are not uniformly distributed but concentrated in certain areas. The Attention Mechanism enhances these important regions while suppressing less significant areas, enabling effective feature extraction. This approach has significantly improved deep learning model performance, particularly in computer vision and natural language processing fields. Attention Mechanisms are primarily categorized into two types: Channel Attention and Spatial Attention (Woo et al., 2018; Liu et al., 2021). Channel Attention emphasizes important feature maps by utilizing inter-channel relationships, while Spatial Attention highlights crucial areas within an image using spatial information. These two attention mechanisms are often combined, allowing models to effectively capture important information in both channel and spatial dimensions.

In this paper, we apply the Attention Mechanism to license plate images captured in real-world environments, which often suffer from low resolution, blur, and distortion. This application successfully improves image quality and increases recognition accuracy by emphasizing important features. Specifically, we employ a hybrid approach combining Channel Attention and Spatial Attention to selectively emphasize character and number regions on license plates. This method demonstrates more effective performance improvement compared to conventional Super-Resolution techniques.

3. Proposed Method

Cameras used in real-world traffic control systems often suffer from image quality degradation due to various constraints. In particular, image loss occurs due to external factors such as the physical distance between vehicles and cameras, weather changes, and lighting conditions. The limited resolution of cameras significantly affects the accuracy of detecting numbers and characters on vehicle license plates. These issues are major factors that degrade the overall performance of traffic management systems. Previously, to address these problems, Super-Resolution using Interpolation techniques was employed. Traditional methods like Bicubic or Bilinear Interpolation were used in large-scale systems such as traffic management due to their advantages of requiring far fewer resources and simpler implementation compared to Convolutional Neural Network-based Super-Resolution methods, thus providing high real-time performance. However, when enlarging the acquired images at high magnification, these methods had limitations in restoring the fine details within license plates, resulting in decreased character and number recognition rates. In this paper, we propose a novel Single Image Super-Resolution (SISR) model that combines a Multi-Scale structure with an Attention Mechanism to address these problems. This model aims to restore low-resolution and degraded images to high-resolution while simultaneously improving image quality. The proposed SISR model incorporates a Multi-Scale Branch to effectively extract features of various sizes, enabling the capture of diverse details in license plate images. Additionally, we employ a combination of Channel Attention and Spatial Attention to emphasize important features of vehicle license plates. Our approach is designed to overcome the challenges posed by real-world conditions in traffic surveillance. By utilizing the Multi-Scale structure, the model can adapt to varying levels of detail in license plate images, from fine textures to larger structural elements. The integration of Attention Mechanisms allows the model to focus on the most crucial areas of the license plate, potentially improving the accuracy of character recognition even in suboptimal imaging conditions. This hybrid approach of Multi-Scale feature extraction and Attention Mechanisms offers a robust solution for enhancing the quality and resolution of license plate images. It addresses not only the issue of low resolution but also compensates for image degradation caused by environmental factors. By doing so, our model aims to significantly improve the performance of license plate recognition systems, contributing to more efficient and accurate traffic management.

Figure 4 illustrates the overall structure of our proposed model, showcasing how the Multi-Scale Branch and Attention Mechanisms are integrated to achieve superior image restoration and enhancement.

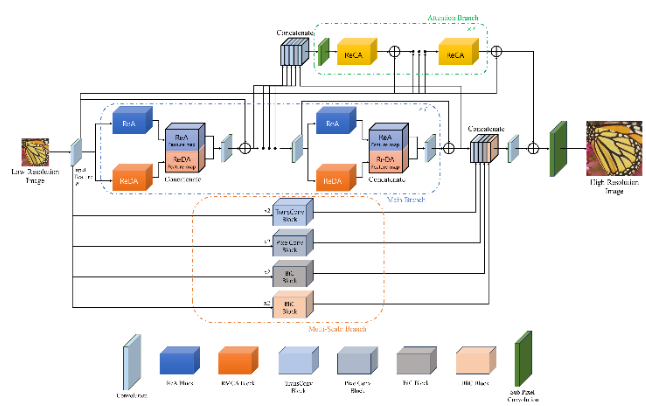

Figure 1: The overall structure of the Proposed Method

Figure 1 illustrates the overall structure of the Single Image Super-Resolution (SISR) network proposed in this study. First, the Main Branch, represented by the blue box in the center, performs learning without changing the size of the feature map of the input Low-Resolution Image. The input feature map is given as input to both the Recurrent Attention

Block (ReA) and the Recurrent Dilated Attention Block (ReDA). Each block extracts different feature maps and performs feature emphasis through Attention. In the case of the ReA Block, features are extracted through Convolution, and emphasis on the feature map is performed repeatedly using Attention multiple times.

Next, the Multi-Scale Branch, represented by the orange box, performs learning by transforming the size of the feature map generated from the Low-Resolution input. The feature map size is expanded to twice its original size using Up-sampling techniques such as Bicubic/Bilinear Interpolation (BiC, BilC), Transposed Convolution (TranConv), and Pixel-Shuffle (PixelConv). After this, feature extraction and emphasis are performed using Convolution and Attention. The resulting features are then combined with the feature maps extracted from the Main Branch, processed with the feature maps generated in the Attention Branch, and finally used to restore the High-Resolution image.

Lastly, the Attention Branch performs learning using the feature maps generated at each stage in the Main Branch as input. In the Main Branch, convolution operations are applied to the feature maps generated through the ReA and ReDA Blocks, and the resulting feature maps are passed as input to the Attention Branch. These received feature maps undergo convolution operations through Recurrent Convolution Attention (ReCA), followed by channel and spatial emphasis to highlight important features. The final output of the network is produced by combining and convolving the feature maps generated through the Main Branch and the Multi-Scale Branch, performing additional operations on the feature maps from the Attention Branch, and finally restoring the high-resolution image through Sub-Pixel Convolution.

Through this complex structure, the proposed model effectively utilizes information from various scales, preserving the detailed information of the low-resolution image while significantly improving overall image quality. In particular, the combination of the Multi-Scale approach and Attention Mechanism effectively restored image features of various sizes and shapes.

4. Experimental Result

In this paper, to validate the performance of the proposed vehicle license plate recognition method, we evaluated the image quality of the high-resolution images restored using our proposed method. Additionally, we performed restoration of vehicle license plates and assessed their recognition accuracy. The proposed model was trained using the DIV2K Dataset and a vehicle license plate OCR Dataset. First, the DIV2K dataset, widely used in SISR models, consists of 800 training images, 100 validation images, and 100 test images. Furthermore, the vehicle license plate OCR dataset provided by AI-Hub includes 80,000 images of various types and colors of license plates, along with corresponding JSON format labeling data. The model was trained using these two datasets. For performance evaluation, we used the Set5, Set14, B100, and Urban100 datasets commonly used in SISR models, as well as 10,000 images from a vehicle license plate validation dataset for testing. For license plate recognition, we used SIGA, a Deep Learning-based text detection model. Low-Resolution images used in training were generated by down-sampling High-Resolution images using Bicubic Interpolation. To evaluate the proposed model's performance, we utilized two main metrics: Peak Signal-to-Noise Ratio (PSNR), which assesses image quality loss, and Structural Similarity Index Measure (SSIM), which evaluates visual quality differences and similarities. Figure 2 shows the result images restored using our proposed model on the Set5, Set14, BSD100, and Urban100 Datasets.

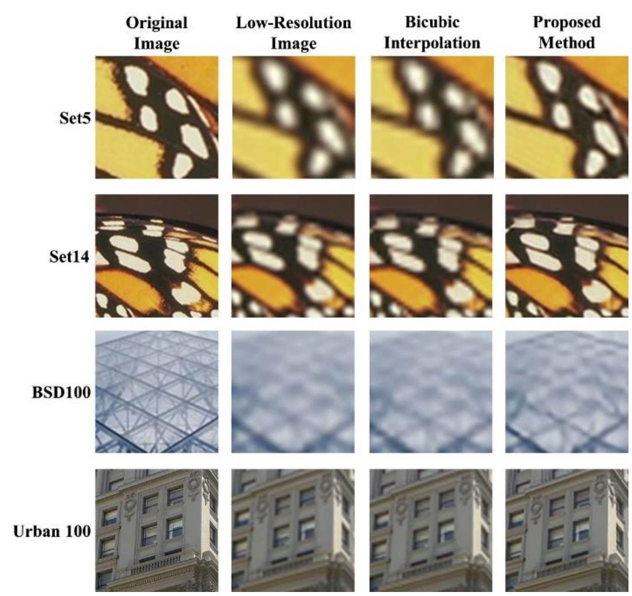

Figure 2: Comparison of Super-Resolution (SR) Benchmark Datasets Using the Proposed Model

Figure 2 demonstrates partial results of applying the proposed Super-Resolution method to various test datasets. The low-resolution images show a significant loss of detail, including edge definition and texture information, compared to the original images. To restore these low-resolution images, we employed both Bicubic Interpolation and our proposed method. While Bicubic Interpolation shows slight improvements in edge and texture information compared to the low-resolution images, it still falls considerably short of the original image quality. In contrast, our proposed method demonstrates remarkable enhancements. In the butterfly wing images from Set5 and Set14, the wing patterns are significantly improved through the restoration of edge and texture information. The BSD100 dataset shows grid patterns that are much sharper than those produced by Bicubic Interpolation. In the Urban100 dataset, architectural details and window edges are substantially enhanced. These results confirm that our proposed method outperforms Bicubic Interpolation in restoring fine details such as edges and textures across various image types, producing results that more closely resemble the original high-resolution images. Figure 3 presents the restoration results of applying our proposed model to a dataset of Korean vehicle license plates.

Table 1: Comparison of Recognition Confidence Scores for License Plate Image Restoration Methods

Plate Number	Confidence Score		
	Original Image	Bicubic Interpolation	Proposed Method
20호 1360	0.9646	0.8445	0.9830
58서 5717	0.9945	0.8758	0.9959
21오 6405	0.8996	0.4214	0.9531
19무 5275	0.7297	0.5323	0.7286
66호 1764	0.6763	0.1714	0.4526
70로 9174	0.5221	0.3765	0.6389

Figure 3: Comparison of License Plate Benchmark Datasets Using the Proposed Model

Figure 3 shows a portion of the results from performing restoration on a vehicle license plate dataset using the proposed method. It can be observed that the Proposed Method significantly improves the overall sharpness of characters and numbers compared to restoration using Bicubic Interpolation. In particular, for the cases of '21오 6405' and '66호 1764', the improvement is notably superior to Bicubic Interpolation. In each result, it can be seen that the outlines of characters and numbers are restored, with the restoration of the Korean characters '오' and '호' showing particularly remarkable improvement over Bicubic Interpolation. This enhancement greatly improves readability by more accurately reproducing the complex structure of Korean characters. Furthermore, it can be observed that noise has been reduced across the entire image, and there is also an improvement in the contrast between the background color of the license plate and the color of the text. Table 1 presents the confidence scores for license plate recognition corresponding to the images shown in Figure 3.

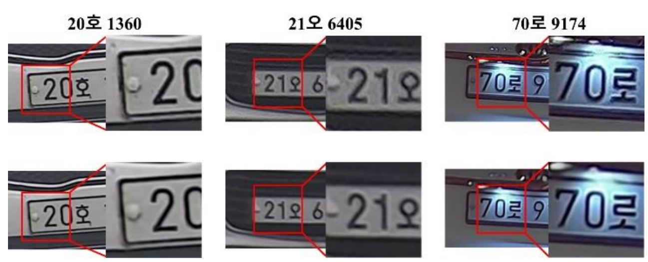

Figure 4: Comparison of Noise Between Results Using the Proposed Model and the Original Image

Table 1 shows the accuracy of character and number recognition within license plates after restoring the vehicle license plate dataset using the proposed method. The experimental results demonstrate that the proposed method outperforms both the Original Image and Bicubic Interpolation in most license plate restorations. Notably, for some license plates (eg., “20호 1360”, “21오 6405”, “70로 9174”), the proposed method shows higher confidence scores than the Original Image. This is due to the noise reduction achieved through the restoration process using the proposed method. Figure 4 illustrates a comparison of noise between the Original Image and the result image. Table 2 presents a comparison of performance between the proposed model and the conventional super-resolution method using Bicubic Interpolation, evaluated using PSNR and SSIM metrics.

Table 2: Comparison of PSNR and SSIM Between Bicubic Interpolation and the Proposed Model

Dataset	Scale	Bicubic Interpolation		Proposed Method	
		PSNR	SSIM	PSNR	SSIM
Set5	X2	33.66	0.9299	38.21	0.9612
	X3	30.39	0.8682	34.76	0.9291
	X4	28.42	0.8104	32.64	0.8988
Set14	X2	30.24	0.8688	34.07	0.9205
	X3	27.55	0.7742	30.68	0.8482
	X4	26.00	0.7027	28.98	0.7899
B100	X2	29.56	0.8431	32.41	0.9018
	X3	27.21	0.7385	29.36	0.8105
	X4	25.96	0.6675	27.77	0.7439
Urban100	X2	26.88	0.8403	33.13	0.9363
	X3	24.46	0.7349	29.04	0.8684
	X4	23.14	0.6577	26.88	0.8086
South Korean License Plate	X2	25.17	0.8199	31.17	0.9698
	X3	24.76	0.7441	29.14	0.9110
	X4	23.04	0.7126	27.28	0.8765

Table 2 compares the PSNR and SSIM performance of Bicubic Interpolation and our proposed method across various datasets. Set5, comprised of 5 images with relatively simple patterns, demonstrates the highest PSNR and SSIM performance among all datasets. For instance, at 2x upscaling, our proposed method achieved 38.21dB PSNR and 0.9612 SSIM. Set14 consists of 14 images from diverse environments, featuring more complex patterns than Set5, resulting in slightly lower performance. At 2x upscaling, our method recorded 34.07dB PSNR and 0.9205 SSIM. B100 contains 100 images with even more intricate patterns, including various subjects such as plants, animals, and people. For this dataset, our method achieved 32.41dB PSNR and 0.9018 SSIM at 2x upscaling. Urban100 is a dataset of 100 architectural images, characterized by complex structures and linear features. Our proposed method exhibited excellent performance on this dataset, with 33.13dB PSNR and 0.9363 SSIM at 2x upscaling. The South Korean License Plate dataset comprises 10,000 diverse Korean vehicle license plate images. Our method achieved 31.17dB PSNR and 0.9698 SSIM at 2x upscaling on this dataset. Across all datasets and upscaling factors (2x, 3x, 4x), our proposed SISR model consistently outperformed Bicubic Interpolation in both PSNR and SSIM metrics. Notably, the performance advantage of our method became more pronounced as the upscaling factor increased. For example, at 4x upscaling on the Urban100 dataset, our method surpassed Bicubic Interpolation by a significant margin of 3.74dB in PSNR and 0.1509 in SSIM.

5. Conclusions

In this paper, we proposed an innovative Super-Resolution model utilizing Multi-Scale and Attention Mechanisms to address the low recognition rate of vehicle license plates due to low-resolution images captured by fixed cameras. Traditional vehicle license plate recognition systems have faced various constraints. Significant image quality loss occurred due to factors such as the physical distance between fixed traffic monitoring cameras and vehicles, vehicle movement, and external environmental conditions like weather and lighting. Moreover, camera performance limitations resulted in the acquisition of only low-resolution images, persistently leading to reduced license plate recognition rates.

To tackle these issues, our research adopted an approach combining Multi-Scale and Attention Mechanisms. Specifically, we designed a network with branch configurations and parallel structures to extract feature maps of various sizes and emphasize crucial features. This structure enabled the effective extraction of information at different scales during the learning process and allowed for the emphasis of key license plate features.

Experimental results demonstrated that our proposed model significantly improved vehicle license plate recognition rates compared to traditional super-resolution restoration methods using Bicubic Interpolation.

Reference

Arica, N., & Yarman-Vural, F. T. (2002). Optical character recognition for cursive handwriting. *IEEE transactions on pattern analysis and machine intelligence*, *24*(6), 801-813.

Bugeja, M., Dingli, A., Attard, M., & Seychell, D. (2020). Comparison of vehicle detection techniques applied to IP camera video feeds for use in intelligent transport systems. *Transportation Research Procedia*, *45*, 971-978.

Cai, W., & Wei, Z. (2020). Remote sensing image classification based on a cross-attention mechanism and graph convolution. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1-5.

Chen, Y., Peng, G., Zhu, Z., & Li, S. (2020). A novel deep learning method based on attention mechanism for bearing remaining useful life prediction. *Applied Soft Computing*, *86*, 105919.

Cui, X., Wang, Q., Dai, J., Xue, Y., & Duan, Y. (2021). Intelligent crack detection based on attention mechanism in convolution neural network. *Advances in Structural Engineering*, *24*(9), 1859-1868.

Dong, C., Loy, C. C., He, K., & Tang, X. (2014). Learning a deep convolutional network for image super-resolution. In Computer Vision–ECCV 2014: *13th European Conference, Zurich, Switzerland, September* 6-12, 2014, Proceedings, Part IV 13 (pp. 184-199). Springer International Publishing.

Drobac, S., & Lindén, K. (2020). Optical character recognition with neural networks and post-correction with finite state methods. *International Journal on Document Analysis and Recognition (IJDAR)*, *23*(4), 279-295.

Dui, H., Zhang, S., Liu, M., Dong, X., & Bai, G. (2024). IoT-enabled real-time traffic monitoring and control management for intelligent transportation systems. *IEEE Internet of Things Journal.*

El Bahi, H., & Zatni, A. (2019). Text recognition in document images obtained by a smartphone based on deep convolutional and recurrent neural network. *Multimedia tools and applications*, *78*(18), 26453-26481.

Guerrero-Ibañez, J., Contreras-Castillo, J., & Zeadally, S. (2021). Deep learning support for intelligent transportation systems. *Transactions on Emerging Telecommunications Technologies*, *32*(3), e4169.

Hassan, T., El-Mowafy, A., & Wang, K. (2021). A review of system integration and current integrity monitoring methods for positioning in intelligent transport systems. *IET Intelligent Transport Systems, 15*(1), 43-60.

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., ... & Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4681-4690).

Lieskovská, E., Jakubec, M., Jarina, R., & Chmulík, M. (2021). A review on speech emotion recognition using deep learning and attention mechanism. *Electronics*, *10*(10), 1163.

Liu, Y., Shao, Z., & Hoffmann, N. (2021). Global attention mechanism: Retain information to enhance channel-spatial interactions. *arXiv preprint arXiv*:2112.05561.

Memon, J., Sami, M., Khan, R. A., & Uddin, M. (2020). Handwritten optical character recognition (OCR): A comprehensive systematic literature review (SLR). *IEEE access*, *8*, 142642-142668.

Ngeni, F. C., Mwakalonge, J. L., Comert, G., Siuhi, S., & Vaidyan, V. (2022). Monitoring of Illegal Removal of Road Barricades Using Intelligent Transportation Systems in Connected and Non-Connected Environments.

Qiu, L., Zhang, D., Tian, Y., & Al-Nabhan, N. (2021). Deep learning-based algorithm for vehicle detection in intelligent transportation systems. *The Journal of Supercomputing, 77*(10), 11083-11098.

Qiu, M., Christopher, L., Chien, S. Y. P., & Chen, Y. (2024). Intelligent Highway Adaptive Lane Learning System in Multiple ROIs of Surveillance Camera Video. *IEEE Transactions on Intelligent Transportation Systems.*

Saharia, C., Ho, J., Chan, W., Salimans, T., Fleet, D. J., & Norouzi, M. (2022). Image super-resolution via iterative refinement. *IEEE transactions on pattern analysis and machine intelligence, 45*(4), 4713-4726.

Salvetti, F., Mazzia, V., Khaliq, A., & Chiaberge, M. (2020). Multi-image super resolution of remotely sensed images using residual attention deep neural networks. *Remote Sensing*, *12*(14), 2207.

Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 3-19).

Yang, J., Wright, J., Huang, T. S., & Ma, Y. (2010). Image super-resolution via sparse representation. *IEEE transactions on image processing, 19*(11), 2861-2873.